Intro to AI philosophy

Strong AI:

If a machine approaches or supersedes human intelligence, if it can do typically human tasks, if it can apply a wide range of background knowledge and has some degree of self-consciousness

Weak AI:

The use of software to study or accomplish specific problem solving or reasoning tasks that do not encompass (or in some cases, are completely outside of) the full range of human cognitive abilities.

History

- The field of artificial intelligence research was founded at a conference on the campus of Dartmouth College in the summer of 1956.

- Those who attended would become the leaders of AI research for decades.

- Many of them predicted that a machine as intelligent as a human being would exist in no more than a generation and they were given millions of dollars to make this vision come true.

What really happened during 60s:

- Simple robot manipulators understanding very limited natural language, able to stack boxes etc.
- Software systems able to perform very simple logical derivations
- First chess and checkers programs
- Crude machine translation

ALPAC committee

ALPAC (Automatic Language Processing Advisory Committee) was a committee of seven scientists led by John R. Pierce, established in 1964 by the U. S. Government in order to evaluate the progress in computational linguistics in general and machine translation in particular.

Its report, issued in 1966, gained notoriety for being very skeptical of research done in machine translation so far, and emphasizing the need for basic research in computational linguistics; this eventually caused the U. S. Government to reduce its funding of the topic dramatically.

„AI freeze“:

Eventually it became obvious that they had grossly underestimated the difficulty of the project. In 1973, in response to the criticism of Sir James Lighthill and ongoing pressure from congress, the U.S. and British Governments stopped funding undirected research into artificial intelligence.

Seven years later, a visionary initiative by the Japanese Government inspired governments and industry to provide AI with billions of dollars (Japanese „5-th generation“ project) , but by the late 80s the investors became disillusioned and withdrew funding again.

Crucial to understand:

- People do not sense, understand or reflect on how their brains really work
- The seemingly „trivial“ everyday activities are actually extremely complex.
- „Simple AI“ is impossible. There is no „clever trick“ to be found.

AI as a part of „cognitive science“

Artificial Intelligence belongs to the family of cognitive science, a border science including following aspects:

- cognitive psychology
- neurology
- philosophy of mind
- ethology
- logic
- evolutionary computing
- sociology and other

Broad schools of thought in AI

Symbolic functionalism

- intelligence represented in symbols and mutual manipulations
- formal models of reasoning, knowledge based

Connectionism

- inspired by natural processes
- emergency of intelligent behaviour

Robotics functionalism

- computational implementation of behaviourism

Symbolic functionalism

Functional Hypothesis

Intelligent behaviour of a system is achieved through interaction among several system's components of different functionality

Physical Symbol Systems Hypothesis

Physical Symbol Systems is an inevitable and satisfactory means for implementation of Intelligent behaviour

Example project: CYC by Doug Lenat

Connectionism

- Essence of intelligence is rooted in static interconnection of huge number of simple processing units and in parallel behaviour of the system as a whole.
- Inspiration from the human brain operations that is a unarguably a medium of intelligent behaviour
- Neuron as a this simple processing unit gets excited according to weighted sum of incoming precepts
- Artificial Neural Networks belong to the field of connectionism
- Distributed Artificial Intelligence – resembles features of connectionism.
- **Example projects: learning systems a la google**

Robotic functionalism

- It is based on computational implementation of the behaviourism as a psychological school of thought
- While symbolic functionalism investigates reasoning background behind intelligent behaviour, robotic functionalism rather explores and simulates outer aspects of intelligent behaviour
- Though the intelligence here is also rooted in interaction among systems components, unlike connectionism they are heterogeneous and hierarchically structured
- Intelligent behaviour is viewed as a sensible interaction among

  agent × environment × task

- **Example project: Rodney Brooks's tiny robots**

IBM Deep Blue vs Garry Kasparov 1997: 3,5 x 2,5

Not AI in the philosophical sense?

Daniel Wolpert claims:

- The brain evolved, not to think or feel, but to control movement

- Take that game of chess as an excellent demonstration of the brain at work; however, if you focus less on the game and more on the movement of the pieces, you'll see something else extraordinary: While computers are capable of sometimes beating the world's best (human) chess masters, "**when it comes to dexterity, a five-year-old child could beat any machine being made"**.

- The brain in its motor-control role is a probabilistic computer, well suited to an uncertain, statistically noisy world. At great speed, Wolpert explained, it compares current sensory data to past experience and weighs probabilities to direct muscles – hundreds of which can be acting at one time -- toward a desired action. And it is able to make decisions long before it gets feedback to see if it has sent the right command. The body "has very long time delays," Wolpert discussed with the Newsletter. "When you send a command to muscles, it takes a very long time to have an effect."

- ttp://www.ted.com/talks/daniel_wolpert_the_real_reason_for_brains

Embodied mind approach

- **Embodied mind thesis** holds that the nature of the human mind is largely determined by the form of the human body.

- George Lakoff argues that **all cognition is based on knowledge that comes from the body** and that other domains are mapped onto our embodied knowledge using a combination of conceptual metaphor, image schema and prototypes.

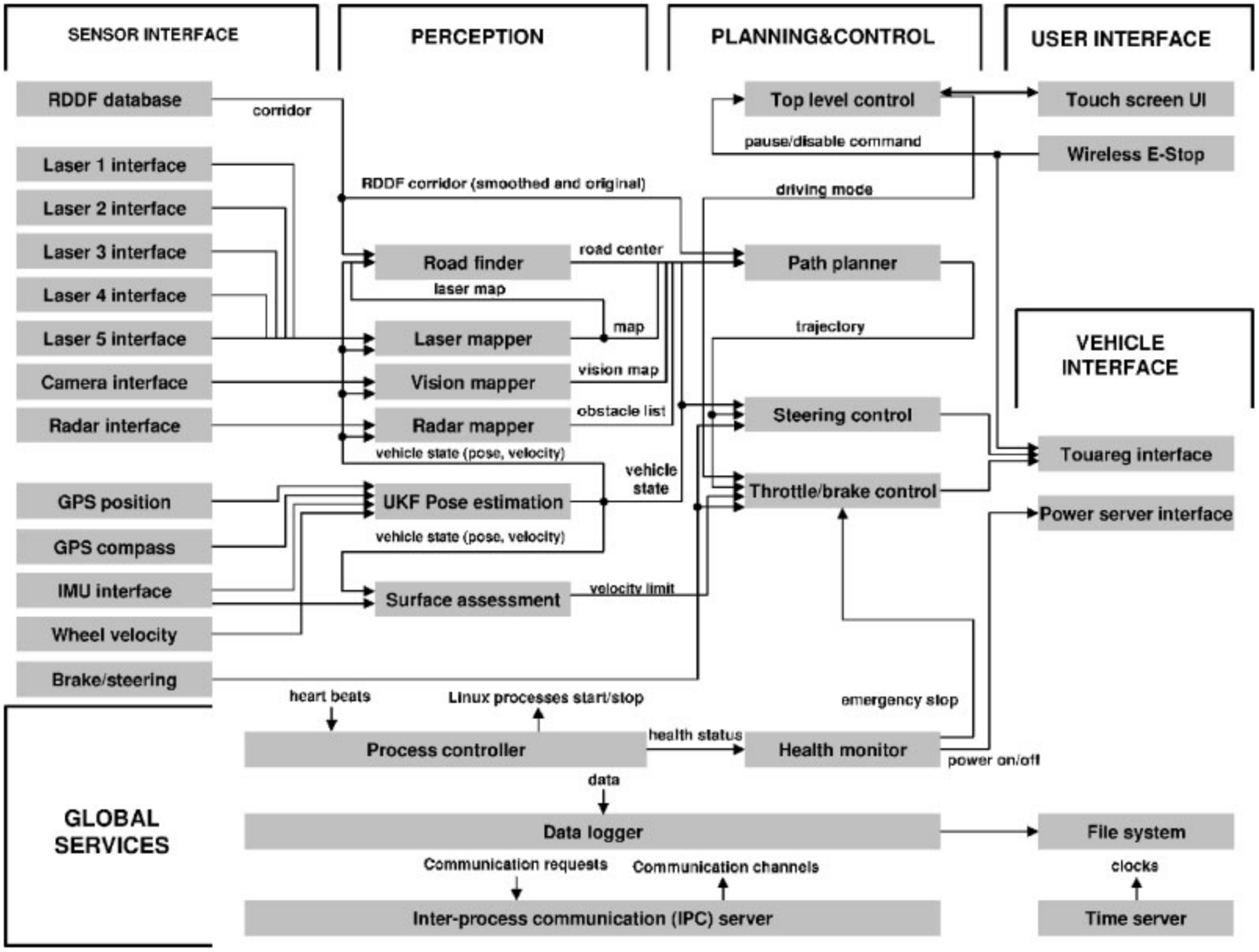
- Example: **love is a journey metaphor.** It is used in such expression as: "we arrived at a crossroads," "we parted ways", "we hit the rocks" (as in a sea journey), "she's in the driver's seat", or, simply, "we're together“.

- Thomas Metzinger claims: no such thing as a self exists. **The conscious self is the content of a model created by our brain—an internal image**, but one we cannot experience as an image. Everything we experience is “a virtual self in a virtual reality.”

Google autonomous car:

Based on vehicle Stanley which won the 2005 DARPA Grand Challenge.

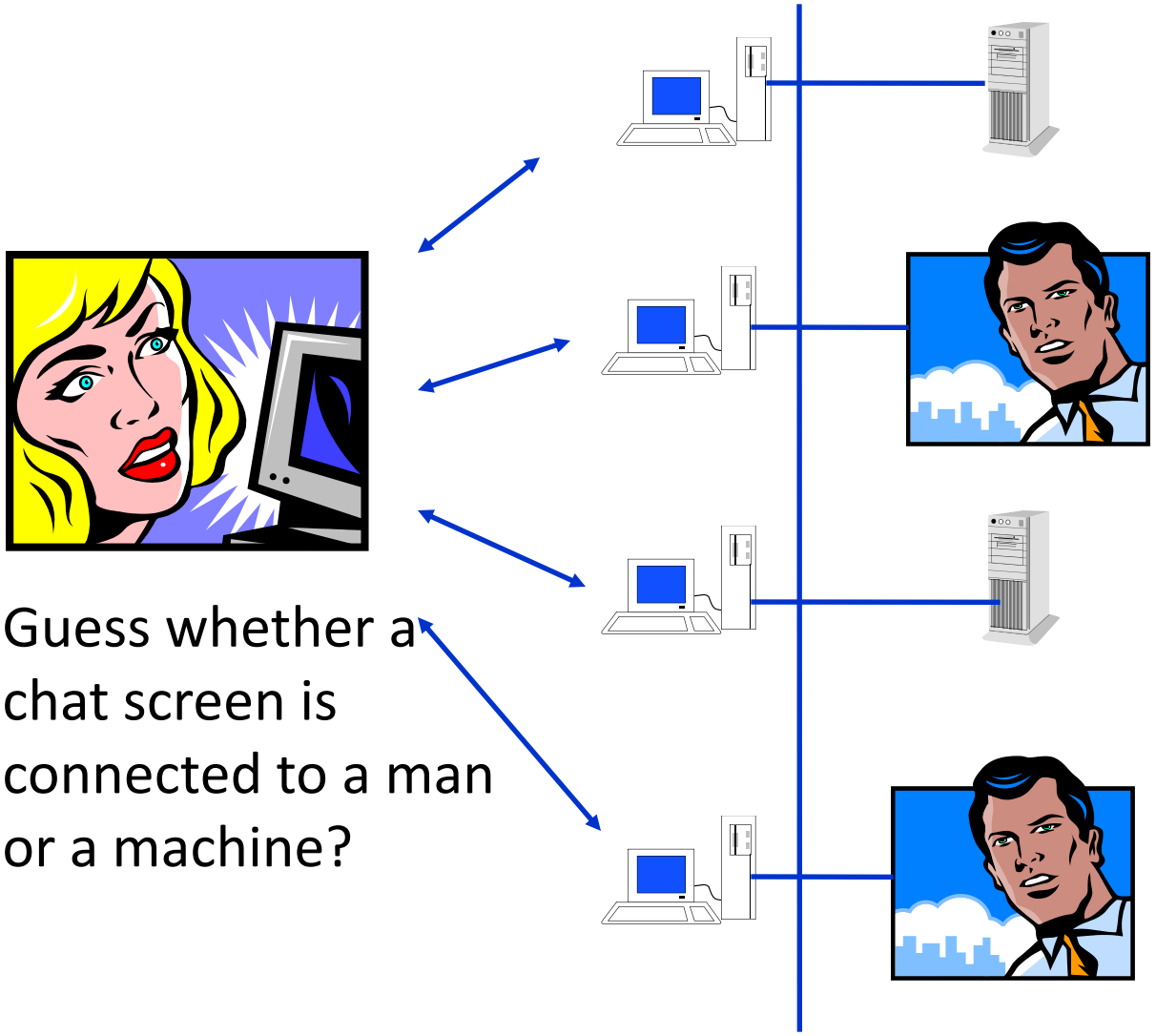
In August 2012, the team announced that they have completed over 300,000 autonomous-driving miles (500 000 km) accident-free


SENSOR INTERFACE
PERCEPTION
PLANNING&CONTROL
USER INTERFACE
RDDF database
corridor
Laser 1 interface
Laser 2 interface
Laser 3 interface
Laser 4 interface
Laser 5 interface
Camera interface
Radar interface
GPS position
GPS compass
IMU interface
Wheel velocity
Brake/steering
RDDF corridor (smoothed and original)
Road finder
road center
laser map
Laser mapper
map
Vision mapper
vision map
Radar mapper
obstacle list
vehicle state (pose, velocity)
UKF Pose estimation
vehicle state
vehicle state (pose, velocity)
Surface assessment
velocity limit
Top level control
pause/disable command
driving mode
Path planner
trajectory
Steering control
Throttle/brake control
Touch screen UI
Wireless E-Stop
VEHICLE INTERFACE
Touareg interface
Power server interface
heart beats
Linux processes start/stop
emergency stop
Process controller
health status
Health monitor
power on/off
data
GLOBAL SERVICES
Data logger
File system
Communication requests
Communication channels
clocks
Inter-process communication (IPC) server
Time server

Popular philosophical questions

- Is it possible for a man to create an intelligent system at all?
- Is it possible to distinguish a „seemingly“ intelligent system from a „really intelligent“ system.
- What is the „essence“ of intelligence?
- What will happen if an intelligent system is created?

Turing test: philosophical suggestion from 30-s:



Guess whether a chat screen is connected to a man or a machine?

Turing suggested:
If humans cannot reliably distinguish a program from a man, then the program is actually intelligent

Philosophy and understanding

We distinguish weak property of a system (its specificity) and strong property (its generality) :

- **weak understanding** (Turing): such an understanding if inputs so that it makes the system to react intelligently (as human)

- **strong understanding** (Brentano): such that the system feels the feeling of understanding as human do

Strong and weak artificial intelligence.

Functionalism

- Claim: There is a level of abstraction below which the specific implementation doesn't matter
- A computer and a brain could both implement this same level of abstraction, and therefore be "isomorphic"
- Strong AI claims that this level exists: if human brains are "thinking", then so are properly-programmed computers.

Continue there:

- **What wikipedia says:**

  http://en.wikipedia.org/wiki/Philosophy_of_artificial_intelligence

  http://en.wikipedia.org/wiki/List_of_artificial_intelligence_projects

- **What reddit says:**

  http://www.reddit.com/r/PhilosophyOfCS

- **What Plato says:**

  http://plato.stanford.edu/entries/computer-science/

- **The brain of C elegans:**

  http://www.artificialbrains.com/openworm

- **Wolpert: the brain evolved, not to think or feel, but to control movement:**

  http://www.ted.com/talks/daniel_wolpert_the_real_reason_for_brains